```
GCCCCAGGGCCTGGAGAGGTCTGAAGAAACCTGGGAGCCAGCAGCCCGGGGCTCCACTCTGGGTTCTGAAAGCCCATTC  79

CCTGCTCTGCGGCTCCTCCCACCCCACCTCTTCTCAGCCTTGCAGCTCAAGGGTTGATCTCAGGAGTCCAGGACCCAGG  158

AGAGGGAAGAATCTGAGGAACACAGAACAGTGAGCGTTGCCCACACCCCATCTCCCGTCACCACATCTCCCCTCACCCT  237

CACCCTCCCTGCCTGGCCCTGGACCCCATCCCAGGACCTCCCTATCAGCTGACTTCTTCCAGTGTCTTGCAGGCCCCTC  316

TGGGCTCCTCCCTCCCCTGGCTTTTCCTACCACTCCCCCTCTATCGGCGTCTATCTGTAGGTGCCCTGGGATTTATAAA  395

ACTGGGTTCCGAATGCTGAATAAGAGACGGTAAGAGCCAAGGCAAAGGACAGCACTGTTCTCTGCCTGCCTGATACCCT  474

                                                    M   A   G   G   A   W   G        7
CACCACCTGGGAACATCCCCCAGACACCCTCTTAACTCCGGGACAGAG ATG GCT GGC GGA GCC TGG GGC       543

 R   L   A   C   Y   L   E   F   L   K   K   E   E   L   K   E   F   Q   L   L     27
CGC CTG GCC TGT TAC TTG GAG TTC CTG AAG AAG GAG GAG CTG AAG GAG TTC CAG CTT CTG    603

 L   A   N   K   A   H   S   R   S   S   S   G   E   T   P   A   Q   P   E   K     47
CTC GCC AAT AAA GCG CAC TCC AGG AGC TCT TCG GGT GAG ACA CCC GCT CAG CCA GAG AAG    663

 T   S   G   M   E   V   A   S   Y   L   V   A   Q   Y   G   E   Q   R   A   W     67
ACG AGT GGC ATG GAG GTG GCC TCG TAC CTG GTG GCT CAG TAT GGG GAG CAG CGG GCC TGG    723

 D   L   A   L   H   T   W   E   Q   M   G   L   R   S   L   C   A   Q   A   Q     87
GAC CTA GCC CTC CAT ACC TGG GAG CAG ATG GGG CTG AGG TCA CTG TGC GCC CAA GCC CAG    783

 E   G   A   G   H   S   P   S   F   P   Y   S   P   S   E   P   H   L   G   S    107
GAA GGG GCA GGC CAC TCT CCC TCA TTC CCC TAC AGC CCA AGT GAA CCC CAC CTG GGG TCT    843

 P   S   Q   P   T   S   T   A   V   L   M   P   W   I   H   E   L   P   A   G    127
CCC AGC CAA CCC ACC TCC ACC GCA GTG CTA ATG CCC TGG ATC CAT GAA TTG CCG GCG GGG    903

 C   T   Q   G   S   E   R   R   V   L   R   Q   L   P   D   T   S   G   R   R    147
TGC ACC CAG GGC TCA GAG AGA AGG GTT TTG AGA CAG CTG CCT GAC ACA TCT GGA CGC CGC    963

 W   R   E   I   S   A   S   L   L   Y   Q   A   L   P   S   S   P   D   H   E    167
TGG AGA GAA ATC TCT GCC TCA CTC CTC TAC CAA GCT CTT CCA AGC TCC CCA GAC CAT GAG   1023

 S   P   S   Q   E   S   P   N   A   P   T   S   T   A   V   L   G   S   W   G    187
TCT CCA AGC CAG GAG TCA CCC AAC GCC CCC ACA TCC ACA GCA GTG CTG GGG AGC TGG GGA   1083

 S   P   P   Q   P   S   L   A   P   R   E   Q   E   A   P   G   T   Q   W   P    207
TCC CCA CCT CAG CCC AGC CTA GCA CCC AGA GAG CAG GAG GCT CCT GGG ACC CAA TGG CCT   1143

 L   D   E   T   S   G   I   Y   Y   T   E   I   R   E   R   E   R   E   K   S    227
CTG GAT GAA ACG TCA GGA ATT TAC TAC ACA GAA ATC AGA GAA AGA GAG AGA GAG AAA TCA   1203

 E   K   G   R   P   P   W   A   A   V   V   G   T   P   P   Q   A   H   T   S    247
GAG AAA GGC AGG CCC CCA TGG GCA GCG GTG GTA GGA ACG CCC CCA CAG GCG CAC ACC AGC   1263

 L   Q   P   H   H   H   P   W   E   P   S   V   R   E   S   L   C   S   T   W    267
CTA CAG CCC CAC CAC CAC CCA TGG GAG CCT TCT GTG AGA GAG AGC CTC TGT TCC ACA TGG   1323

 P   W   K   N   E   D   F   N   Q   K   F   T   Q   L   L   L   L   Q   R   P    287
CCC TGG AAA AAT GAG GAT TTT AAC CAA AAA TTC ACA CAG CTG CTA CTT CTA CAA AGA CCT   1383

 H   P   R   S   Q   D   P   L   V   K   R   S   W   P   D   Y   V   E   E   N    307
CAC CCC AGA AGC CAA GAT CCC CTG GTC AAG AGA AGC TGG CCT GAT TAT GTG GAG GAG AAT   1443

 R   G   H   L   I   E   I   R   D   L   F   G   P   G   L   D   T   Q   E   P    327
CGA GGA CAT TTA ATT GAG ATC AGA GAC TTA TTT GGC CCA GGC CTG GAT ACC CAA GAA CCT   1503

 R   I   V   I   L   Q   G   A   A   G   I   G   K   S   T   L   A   R   Q   V    347
CGC ATA GTC ATA CTG CAG GGG GCT GCT GGA ATT GGG AAG TCA ACA CTG GCC AGG CAG GTG   1563
```

FIG. 1A

```
 K   E   A   W   G   R   G   Q   L   Y   G   D   R   F   Q   H   V   F   Y   F   367
AAG GAA GCC TGG GGG AGA GGC CAG CTG TAT GGG GAC CGC TTC CAG CAT GTC TTC TAC TTC  1623

 S   C   R   E   L   A   Q   S   K   V   V   S   L   A   E   L   I   G   K   D   387
AGC TGC AGA GAG CTG GCC CAG TCC AAG GTG GTG AGT CTC GCT GAG CTC ATC GGA AAA GAT  1683

 G   T   A   T   P   A   P   I   R   Q   I   L   S   R   P   E   R   L   L   F   407
GGG ACA GCC ACT CCG GCT CCC ATT AGA CAG ATC CTG TCT AGG CCA GAG CGG CTG CTC TTC  1743

 I   L   D   G   V   D   E   P   G   W   V   L   Q   E   P   S   S   E   L   C   427
ATC CTC GAT GGT GTA GAT GAG CCA GGA TGG GTC TTG CAG GAG CCG AGT TCT GAG CTC TGT  1803

 L   H   W   S   Q   P   Q   P   A   D   A   L   L   G   S   L   L   G   K   T   447
CTG CAC TGG AGC CAG CCA CAG CCG GCG GAT GCA CTG CTG GGC AGT TTG CTG GGG AAA ACT  1863

 I   L   P   E   A   S   F   L   I   T   A   R   T   T   A   L   Q   N   L   I   467
ATA CTT CCC GAG GCA TCC TTC CTG ATC ACG GCT CGG ACC ACA GCT CTG CAG AAC CTC ATT  1923

 P   S   L   E   Q   A   R   W   V   E   V   L   G   F   S   E   S   S   R   K   487
CCT TCT TTG GAG CAG GCA CGT TGG GTA GAG GTC CTG GGG TTC TCT GAG TCC AGC AGG AAG  1983

 E   Y   F   Y   R   Y   F   T   D   E   R   Q   A   I   R   A   F   R   L   V   507
GAA TAT TTC TAC AGA TAT TTC ACA GAT GAA AGG CAA GCA ATT AGA GCC TTT AGG TTG GTC  2043

 K   S   N   K   E   L   W   A   L   C   L   V   P   W   V   S   W   L   A   C   527
AAA TCA AAC AAA GAG CTC TGG GCC CTG TGT CTT GTG CCC TGG GTG TCC TGG CTG GCC TGC  2103

 T   C   L   M   Q   Q   M   K   R   K   E   K   L   T   L   T   S   K   T   T   547
ACT TGC CTG ATG CAG CAG ATG AAG CGG AAG GAA AAA CTC ACA CTG ACT TCC AAG ACC ACC  2163

 T   T   L   C   L   H   Y   L   A   Q   A   L   Q   A   Q   P   L   G   P   Q   567
ACA ACC CTC TGT CTA CAT TAC CTT GCC CAG GCT CTC CAA GCT CAG CCA TTG GGA CCC CAG  2223

 L   R   D   L   C   S   L   A   A   E   G   I   W   Q   K   K   T   L   F   S   587
CTC AGA GAC CTC TGC TCT CTG GCT GCT GAG GGC ATC TGG CAA AAA AAG ACC CTT TTC AGT  2283

 P   D   D   L   R   K   H   G   L   D   G   A   I   I   S   T   F   L   K   M   607
CCA GAT GAC CTC AGG AAG CAT GGG TTA GAT GGG GCC ATC ATC TCC ACC TTC TTG AAG ATG  2343

 G   I   L   Q   E   H   P   I   P   L   S   Y   S   F   I   H   L   C   F   Q   627
GGT ATT CTT CAA GAG CAC CCC ATC CCT CTG AGC TAC AGC TTC ATT CAC CTC TGT TTC CAA  2403

 E   F   F   A   A   M   S   Y   V   L   E   D   E   K   G   R   G   K   H   S   647
GAG TTC TTT GCA GCA ATG TCC TAT GTC TTG GAG GAT GAG AAG GGG AGA GGT AAA CAT TCT  2463

 N   C   I   I   D   L   E   K   T   L   E   A   Y   G   I   H   G   L   F   G   667
AAT TGC ATC ATA GAT TTG GAA AAG ACG CTA GAA GCA TAT GGA ATA CAT GGC CTG TTT GGG  2523

 A   S   T   T   R   F   L   L   G   L   L   S   D   E   G   E   R   E   M   E   687
GCA TCA ACC ACA CGT TTC CTA TTG GGC CTG TTA AGT GAT GAG GGG GAG AGA GAG ATG GAG  2583

 N   I   F   H   C   R   L   S   Q   G   R   N   L   M   Q   W   V   P   S   L   707
AAC ATC TTT CAC TGC CGG CTG TCT CAG GGG AGG AAC CTG ATG CAG TGG GTC CCG TCC CTG  2643

 Q   L   L   L   Q   P   H   S   L   E   S   L   H   C   L   Y   E   T   R   N   727
CAG CTG CTG CTG CAG CCA CAC TCT CTG GAG TCC CTC CAC TGC TTG TAC GAG ACT CGG AAC  2703

 K   T   F   L   T   Q   V   M   A   H   F   E   E   M   G   M   C   V   E   T   747
AAA ACG TTC CTG ACA CAA GTG ATG GCC CAT TTC GAA GAA ATG GGC ATG TGT GTA GAA ACA  2763

 D   M   E   L   L   V   C   T   F   C   I   K   F   S   R   H   V   K   K   L   767
GAC ATG GAG CTC TTA GTG TGC ACT TTC TGC ATT AAA TTC AGC CGC CAC GTG AAG AAG CTT  2823

 Q   L   I   E   G   R   Q   H   R   S   T   W   S   P   T   M   V   V   L   F   787
CAG CTG ATT GAG GGC AGG CAG CAC AGA TCA ACA TGG AGC CCC ACC ATG GTA GTC CTG TTC  2883

 R   W   V   P   V   T   D   A   Y   W   Q   I   L   F   S   V   L   K   V   T   807
```

FIG. 1B

```
AGG TGG GTC CCA GTC ACA GAT GCC TAT TGG CAG ATT CTC TTC TCC GTC CTC AAG GTC ACC  2943

 R   N   L   K   E   L   D   L   S   G   N   S   L   S   H   S   A   V   K   S   827
AGA AAC CTG AAG GAG CTG GAC CTA AGT GGA AAC TCG CTG AGC CAC TCT GCA GTG AAG AGT  3003

 L   C   K   T   L   R   R   P   R   C   L   L   E   T   L   R   L   A   G   C   847
CTT TGT AAG ACC CTG AGA CGC CCT CGC TGC CTC CTG GAG ACC CTG CGG TTG GCT GGC TGT  3063

 G   L   T   A   E   D   C   K   D   L   A   F   G   L   R   A   N   Q   T   L   867
GGC CTC ACA GCT GAG GAC TGC AAG GAC CTT GCC TTT GGG CTG AGA GCC AAC CAG ACC CTG  3123

 T   E   L   D   L   S   F   N   V   L   T   D   A   G   A   K   H   L   C   Q   887
ACC GAG CTG GAC CTG AGC TTC AAT GTG CTC ACG GAT GCT GGA GCC AAA CAC CTT TGC CAG  3183

 R   L   R   Q   P   S   C   K   L   Q   R   L   Q   L   V   S   C   G   L   T   907
AGA CTG AGA CAG CCG AGC TGC AAG CTA CAG CGA CTG CAG CTG GTC AGC TGT GGC CTC ACG  3243

 S   D   C   C   Q   D   L   A   S   V   L   S   A   S   P   S   L   K   E   L   927
TCT GAC TGC TGC CAG GAC CTG GCC TCT GTG CTT AGT GCC AGC CCC AGC CTG AAG GAG CTA  3303

 D   L   Q   Q   N   N   L   D   D   V   G   V   R   L   L   C   E   G   L   R   947
GAC CTG CAG CAG AAC AAC CTG GAT GAC GTT GGC GTG CGA CTG CTC TGT GAG GGG CTC AGG  3363

 H   P   A   C   K   L   I   R   L   G   L   D   Q   T   T   L   S   D   E   M   967
CAT CCT GCC TGC AAA CTC ATA CGC CTG GGG CTG GAC CAG ACA ACT CTG AGT GAT GAG ATG  3423

 R   Q   E   L   R   A   L   E   Q   E   K   P   Q   L   L   I   F   S   R   R   987
AGG CAG GAA CTG AGG GCC CTG GAG CAG GAG AAA CCT CAG CTG CTC ATC TTC AGC AGA CGG  3483

 K   P   S   V   M   T   P   T   E   G   L   D   T   G   E   M   S   N   S   T  1007
AAA CCA AGT GTG ATG ACC CCT ACT GAG GGC CTG GAT ACG GGA GAG ATG AGT AAT AGC ACA  3543

 S   S   L   K   R   Q   R   L   G   S   E   R   A   A   S   H   V   A   Q   A  1027
TCC TCA CTC AAG CGG CAG AGA CTC GGA TCA GAG AGG GCG GCT TCC CAT GTT GCT CAG GCT  3603

 N   L   K   L   L   D   V   S   K   I   F   P   I   A   E   I   A   E   E   S  1047
AAT CTC AAA CTC CTG GAC GTG AGC AAG ATC TTC CCA ATT GCT GAG ATT GCA GAG GAA AGC  3663

 S   P   E   V   V   P   V   E   L   L   C   V   P   S   P   A   S   Q   G   D  1067
TCC CCA GAG GTA GTA CCG GTG GAA CTC TTG TGC GTG CCT TCT CCT GCC TCT CAA GGG GAC  3723

 L   H   T   K   P   L   G   T   D   D   D   F   W   G   P   T   G   P   V   A  1087
CTG CAT ACG AAG CCT TTG GGG ACT GAC GAT GAC TTC TGG GGC CCC ACG GGG CCT GTG GCT  3783

 T   E   V   V   D   K   E   K   N   L   Y   R   V   H   F   P   V   A   G   S  1107
ACT GAG GTA GTT GAC AAA GAA AAG AAC TTG TAC CGA GTT CAC TTC CCT GTA GCT GGC TCC  3843

 Y   R   W   P   N   T   G   L   C   F   V   M   R   E   A   V   T   V   E   I  1127
TAC CGC TGG CCC AAC ACG GGT CTC TGC TTT GTG ATG AGA GAA GCG GTG ACC GTT GAG ATT  3903

 E   F   C   V   W   D   Q   F   L   G   E   I   N   P   Q   H   S   W   M   V  1147
GAA TTC TGT GTG TGG GAC CAG TTC CTG GGT GAG ATC AAC CCA CAG CAC AGC TGG ATG GTG  3963

 A   G   P   L   L   D   I   K   A   E   P   G   A   V   E   A   V   H   L   P  1167
GCA GGG CCT CTG CTG GAC ATC AAG GCT GAG CCT GGA GCT GTG GAA GCT GTG CAC CTC CCT  4023

 H   F   V   A   L   Q   G   G   H   V   D   T   S   L   F   Q   M   A   H   F  1187
CAC TTT GTG GCT CTC CAA GGG GGC CAT GTG GAC ACA TCC CTG TTC CAA ATG GCC CAC TTT  4083

 K   E   E   G   M   L   L   E   K   P   A   R   V   E   L   H   H   I   V   L  1207
AAA GAG GAG GGG ATG CTC CTG GAG AAG CCA GCC AGG GTG GAG CTG CAT CAC ATA GTT CTG  4143

 E   N   P   S   F   S   P   L   G   V   L   L   K   M   I   H   N   A   L   R  1227
GAA AAC CCC AGC TTC TCC CCC TTG GGA GTC CTC CTG AAA ATG ATC CAT AAT GCC CTG CGC  4203

 F   I   P   V   T   S   V   V   L   L   Y   H   R   V   H   P   E   E   V   T  1247
TTC ATT CCC GTC ACC TCT GTG GTG TTG CTT TAC CAC CGC GTC CAT CCT GAG GAA GTC ACC  4263
```

FIG. 1C

```
 F   H   L   Y   L   I   P   S   D   C   S   I   R   K   E   L   E   L   C   Y  1267
TTC CAC CTC TAC CTG ATC CCA AGT GAC TGC TCC ATT CGG AAG GAA CTG GAG CTC TGC TAT  4323

 R   S   P   G   E   D   Q   L   F   S   E   F   Y   V   G   H   L   G   S   G  1287
CGA AGC CCT GGA GAA GAC CAG CTG TTC TCG GAG TTC TAC GTT GGC CAC TTG GGA TCA GGG  4383

 I   R   L   Q   V   K   D   K   K   D   E   T   L   V   W   E   A   L   V   K  1307
ATC AGG CTG CAA GTG AAA GAC AAG AAA GAT GAG ACT CTG GTG TGG GAG GCC TTG GTG AAA  4443

 P   G   D   L   M   P   A   T   T   L   I   P   P   A   R   I   A   V   P   S  1327
CCA GGA GAT CTC ATG CCT GCA ACT ACT CTG ATC CCT CCA GCC CGC ATA GCC GTA CCT TCA  4503

 P   L   D   A   P   Q   L   L   H   F   V   D   Q   Y   R   E   Q   L   I   A  1347
CCT CTG GAT GCC CCG CAG TTG CTG CAC TTT GTG GAC CAG TAT CGA GAG CAG CTG ATA GCC  4563

 R   V   T   S   V   E   V   V   L   D   K   L   H   G   Q   V   L   S   Q   E  1367
CGA GTG ACA TCG GTG GAG GTT GTC TTG GAC AAA CTG CAT GGA CAG GTG CTG AGC CAG GAG  4623

 Q   Y   E   R   V   L   A   E   N   T   R   P   S   Q   M   R   K   L   F   S  1387
CAG TAC GAG AGG GTG CTG GCT GAG AAC ACG AGG CCC AGC CAG ATG CGG AAG CTG TTC AGC  4683

 L   S   Q   S   W   D   R   K   C   K   D   G   L   Y   Q   A   L   K   E   T  1407
TTG AGC CAG TCC TGG GAC CGG AAG TGC AAA GAT GGA CTC TAC CAA GCC CTG AAG GAG ACC  4743

 H   P   H   L   I   M   E   L   W   E   K   G   S   K   K   G   L   L   P   L  1427
CAT CCT CAC CTC ATT ATG GAA CTC TGG GAG AAG GGC AGC AAA AAG GGA CTC CTG CCA CTC  4803

 S   S   *                                                                      1430
AGC AGC TGA                                                                     4812

AGTATCAACACCAGCCCTTGACCCTTGAGTCCTGGCTTTGGCTGACCCTTCTTTGGGTCTCAGTTTCTTTCTCTGCAAA  4891

CAAGTTGCCATCTGGTTTGCCTTCCAGCACTAAAGTAATGGAACTTTGATGATGCCTTTGCTGGGCATTATGTGTCCAT  4970

GCCAGGGATGCCACAGGGGGCCCCAGTCCAGGTGGCCTAACAGCATCTCAGGGAATGTCCATCTGGAGCTGGCAAGACC  5049

CCTGCAGACCTCATAGAGCCTCATCTGGTGGCCACAGCAGCCAAGCCTAGAGCCCTCCGGATCCCATCCAGGCGCAAAG  5128

AGGAATAGGAGGGACATGGAACCATTTGCCTCTGGCTGTGTCACAGGGTGAGCCCCAAAATTGGGGTTCAGCGTGGGAG  5207

GCCACGTGGATTCTTGGCTTTGTACAGGAAGATCTACAAGAGCAAGCCAACAGAGTAAAGTGGAAGGAAGTTTATTCAG  5286

AAAATAAAGGAGTATCACAGCTCTTTTAGAATTTGTCTAGCAGGCTTTCCAGTTTTTACCAGAAAACCCCTATAAATTA  5365

AAAATTTTTTACTTAAATTTAAGAATTAAAAAAATACAAAAAAGAAAAAATGAAAATAAAGGAATAAGAAGTTACCTAC  5444

AAAATTTTTTACTTAAATTTAAGAATTAAAAAAATACAAAAAAGAAAAAATGAAAATAAAGGAATAAGAAGTTACCTAC  5444
```

FIG. 1D

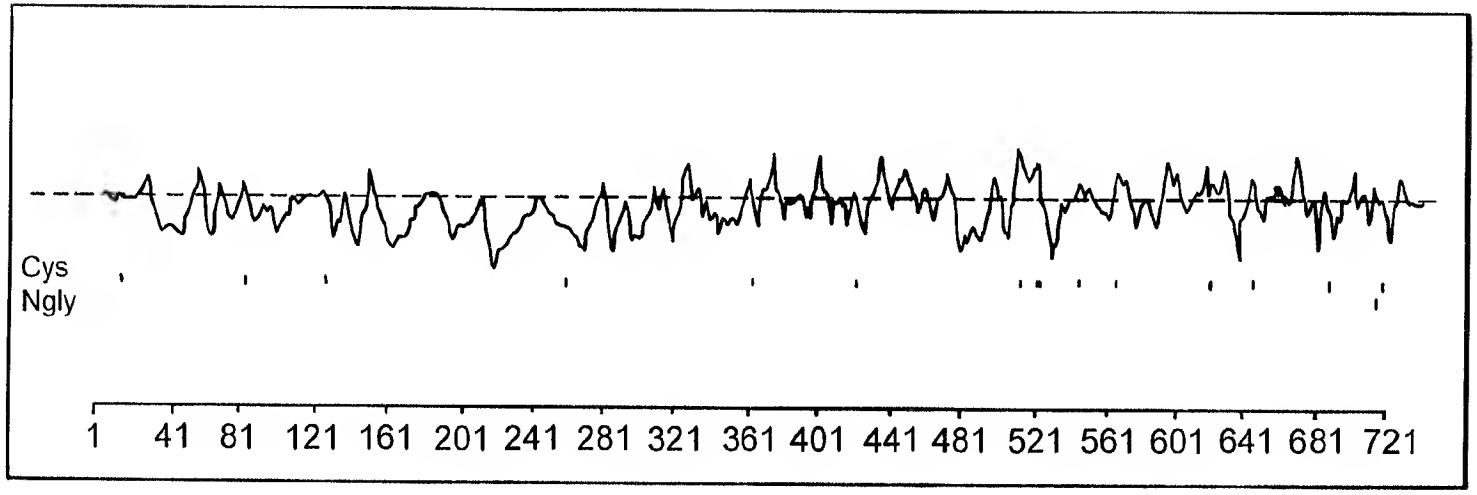
Cys
Ngly
1 41 81 121 161 201 241 281 321 361 401 441 481 521 561 601 641 681 721

FIG. 2

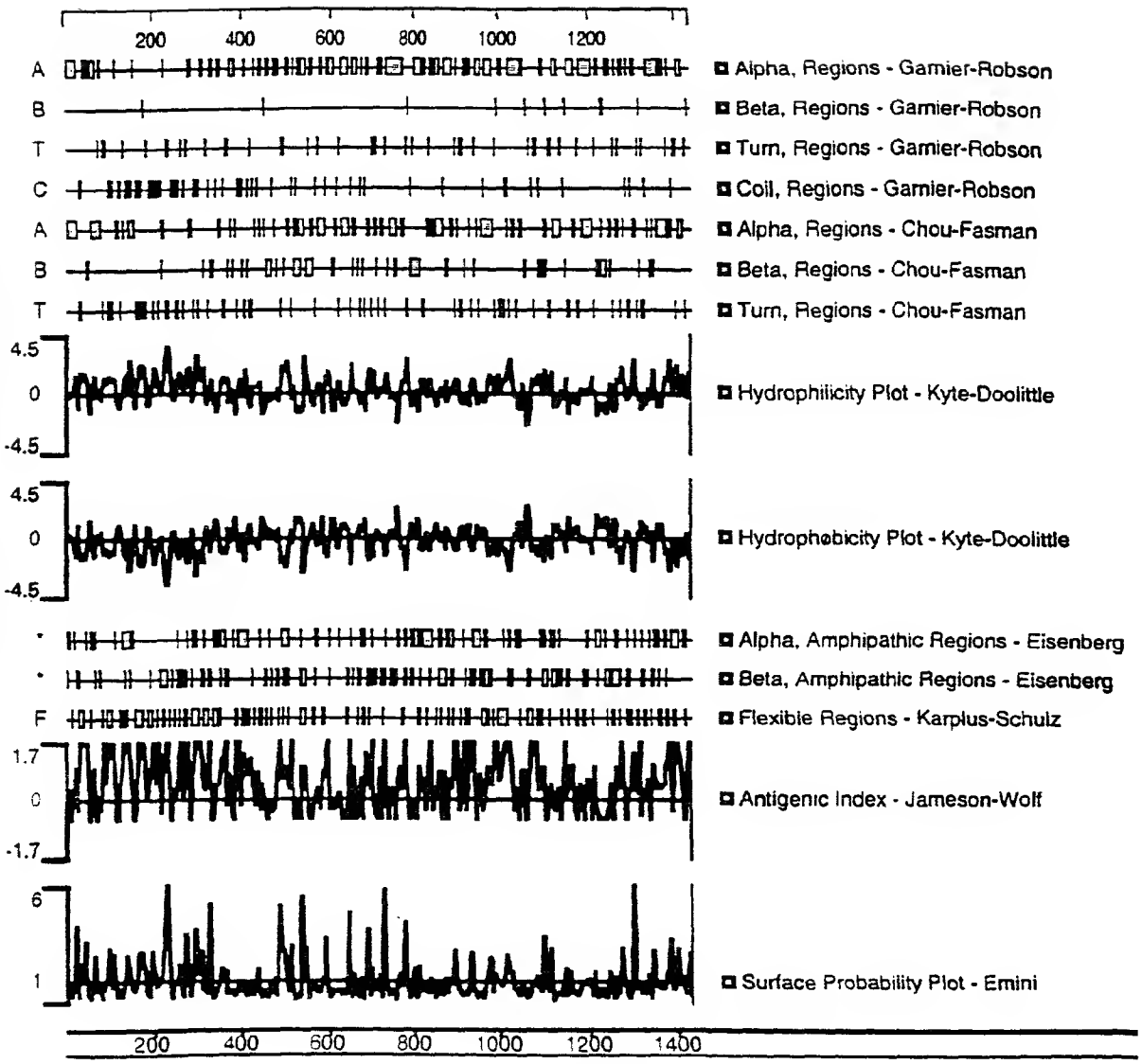
200
400
600
800
1000
1200
A
B
T
C
A
B
T
Alpha, Regions - Garnier-Robson
Beta, Regions - Garnier-Robson
Turn, Regions - Garnier-Robson
Coil, Regions - Garnier-Robson
Alpha, Regions - Chou-Fasman
Beta, Regions - Chou-Fasman
Turn, Regions - Chou-Fasman
4.5
0
-4.5
Hydrophilicity Plot - Kyte-Doolittle
4.5
0
-4.5
Hydrophobicity Plot - Kyte-Doolittle
Alpha, Amphipathic Regions - Eisenberg
Beta, Amphipathic Regions - Eisenberg
F
Flexible Regions - Karplus-Schulz
1.7
0
-1.7
Antigenic Index - Jameson-Wolf
6
1
Surface Probability Plot - Emini
200
400
600
800
1000
1200
1400

FIG. 3

```
CTGGTTCTCAACTTCTTTTGAAATAATGTTCATAGAGAAGGAGGGCTGTCTGAGATTCGAGGGAAACAAGCTCTCAGGA  79

CTTCCGGTCGCCATGATGGCTGTGGGCGGTAAACGCGGTTAGTGCAAGCATCTGGGCCATCTTCAATGGTAAAAAAGAT  158

ACAGTAAAGACATAAATACCACATTTGACAAATGGAAAAAAAGGAGTGTCCAGAAAAGAGTAGCAGCAGTGAGGAAGAG  237

CTGCCGAGACGGGTATACAGGGAGCTACCCTGTGTTTCTGAGACCCTTTGTGACATCTCACATTTTTTCCAAGAAG      313

 M  M  R  Q  R  Q  S  H  Y  C  S  V  L  F  L  S  V  N  Y  L    20
ATG ATG AGA CAG AGG CAG AGC CAT TAT TGT TCC GTG CTG TTC CTG AGT GTC AAC TAT CTG   373

 G  G  T  F  P  G  D  I  C  S  E  E  N  Q  I  V  S  S  Y  A    40
GGG GGG ACA TTC CCA GGA GAC ATT TGC TCA GAA GAG AAT CAA ATA GTT TCC TCT TAT GCT   433

 S  K  V  C  F  E  I  E  E  D  Y  K  N  R  Q  F  L  G  P  E    60
TCT AAA GTC TGT TTT GAG ATC GAA GAA GAT TAT AAA AAT CGT CAG TTT CTG GGG CCT GAA   493

 G  N  V  D  V  E  L  I  D  K  S  T  N  R  Y  S  V  W  F  P    80
GGA AAT GTG GAT GTT GAG TTG ATT GAT AAG AGC ACA AAC AGA TAC AGC GTT TGG TTC CCC   553

 T  A  G  W  Y  L  W  S  A  T  G  L  G  F  L  V  R  D  E  V   100
ACT GCT GGC TGG TAT CTG TGG TCA GCC ACA GGC CTC GGC TTC CTG GTA AGG GAT GAG GTC   613

 T  V  T  I  A  F  G  S  W  S  Q  H  L  A  L  D  L  Q  H  H   120
ACA GTG ACG ATT GCG TTT GGT TCC TGG AGT CAG CAC CTG GCC CTG GAC CTG CAG CAC CAT   673

 E  Q  W  L  V  G  G  P  L  F  D  V  T  A  E  P  E  E  A  V   140
GAA CAG TGG CTG GTG GGC GGC CCC TTG TTT GAT GTC ACT GCA GAG CCA GAG GAG GCT GTC   733

 A  E  I  H  L  P  H  F  I  S  L  Q  G  E  V  D  V  S  W  F   160
GCC GAA ATC CAC CTC CCC CAC TTC ATC TCC CTC CAA GGT GAG GTG GAC GTC TCC TGG TTT   793

 L  V  A  H  F  K  N  E  G  M  V  L  E  H  P  A  R  V  E  P   180
CTC GTT GCC CAT TTT AAG AAT GAA GGG ATG GTC CTG GAG CAT CCA GCC CGG GTG GAG CCT   853

 F  Y  A  V  L  E  S  P  S  F  S  L  M  G  I  L  L  R  I  A   200
TTC TAT GCT GTC CTG GAA AGC CCC AGC TTC TCT CTG ATG GGC ATC CTG CTG CGG ATC GCC   913

 S  G  T  R  L  S  I  P  I  T  S  N  T  L  I  Y  Y  H  P  H   220
AGT GGG ACT CGC CTC TCC ATC CCC ATC ACT TCC AAC ACA TTG ATC TAT TAT CAC CCC CAC   973

 P  E  D  I  K  F  H  L  Y  L  V  P  S  D  A  L  L  T  K  A   240
CCC GAA GAT ATT AAG TTC CAC TTG TAC CTT GTC CCC AGC GAC GCC TTG CTA ACA AAG GCG  1033

 I  D  D  E  E  D  R  F  H  G  V  R  L  Q  T  S  P  P  M  E   260
ATA GAT GAT GAG GAA GAT CGC TTC CAT GGT GTG CGC CTG CAG ACT TCG CCC CCA ATG GAA  1093

 P  L  N  F  G  S  S  Y  I  V  S  N  S  A  N  L  K  V  M  P   280
CCC CTG AAC TTT GGT TCC AGT TAT ATT GTG TCT AAT TCT GCT AAC CTG AAA GTA ATG CCC  1153

 K  E  L  K  L  S  Y  R  S  P  G  E  I  Q  H  F  S  K  F  Y   300
AAG GAG TTG AAA TTG TCC TAC AGG AGC CCT GGA GAA ATT CAG CAC TTC TCA AAA TTC TAT  1213

 A  G  Q  M  K  E  P  I  Q  L  E  I  T  E  K  R  H  G  T  L   320
GCT GGG CAG ATG AAG GAA CCC ATT CAA CTT GAG ATT ACT GAA AAA AGA CAT GGG ACT TTG  1273

 V  W  D  T  E  V  K  P  V  D  L  Q  L  V  A  A  S  A  P  P   340
GTG TGG GAT ACT GAG GTG AAG CCA GTG GAT CTC CAG CTT GTA GCT GCA TCA GCC CCT CCT  1333

 P  F  S  G  A  A  F  V  K  E  N  H  R  Q  L  Q  A  R  M  G   360
CCT TTC TCA GGT GCA GCC TTT GTG AAG GAG AAC CAC CGG CAA CTC CAA GCC AGG ATG GGG  1393

 D  L  K  G  V  L  D  D  L  Q  D  N  E  V  L  T  E  N  E  K   380
GAC CTG AAA GGG GTG CTC GAT GAT CTC CAG GAC AAT GAG GTT CTT ACT GAG AAT GAG AAG  1453

 E  L  V  E  Q  E  K  T  R  Q  S  K  N  E  A  L  L  S  M  V   400
```

FIG. 4A

```
GAG CTG GTG GAG CAG GAA AAG ACA CGG CAG AGC AAG AAT GAG GCC TTG CTG AGC ATG GTG  1513

 E   K   K   G   D   L   A   L   D   V   L   F   R   S   I   S   E   R   D   P   420
GAG AAG AAA GGG GAC CTG GCC CTG GAC GTG CTC TTC AGA AGC ATT AGT GAA AGG GAC CCT  1573

 Y   L   V   S   Y   L   R   Q   Q   N   L   *                                   432
TAC CTC GTG TCC TAT CTT AGA CAG CAG AAT TTG TAA                                   1609

AATGAGTCAGTTAGGTAGTCTGGAAGAGAGAATCCAGCGTTCTCATTGGAAATGGATAAACAGAAATGTGATCATTGAT  1688

TTCAGTGTTCAAGACAGAAGAAGACTGGGTAACATCTATCACACAGGCTTTCAGGACAGACTTGTAACCTGGCATGTAC  1767

CTATTGACTGTATCCTCATGCATTTTCCTCAAGAATGTCTGAAGAAGGTAGTAATATTCCTTTTAAATTTTTTCCAACC  1846

ATTGCTTGATATATCACTATTTTATCCATTGACATGATTCTTGAAGACCCAGGATAAAGGACATCCGGATAGGTGTGTT  1925

TATGAAGGATGGGGCCTGGAAAGGCAACTTTTCCTGATTAATGTGAAAAATAATTCCTATGGACACTCCGTTTGAAGTA  2004

TCACCTTCTCATAACTAAAAGCAGAAAAGCTAACAAAAGCTTCTCAGCTGAGGACACTCAAGGCATACATGATGACAGT  2083

CTTTTTTTTTTTTGTATGTTAGGACTTTAACACTTTATCTATGGCTACTGTTATTAGAACAATGTAAATGTATTTGCTG  2162

AAAGAGAGCACAAAAATGGGAGAAAATGCAAACATGAGCAGAAAATATTTTCCCACTGGTGTGTAGCCTGCTACAAGGA  2241

GTTGTTGGGTTAAATGTTCATGGTCAACTCCAAGGAATACTGAGATGAAATGTGGTAAATCAACTCCACAGAACCACCA  2320

AAAAGAAAATGAGGGTAATTCAGCTTATTCTGAGACAGACATTCCTGGCAATGTACCATACAAAAAATAAGCCAACTCT  2399

GACATTTGGATTCTACCATAGACTCTGTCATTTTGTAGCCATTTCAGCTGTCTTTTGATTAATGTTTTCGTGGCACACA  2478

TATTTCCATCCTTTTATGTTTAATCTGTTTAAAACAAGTTCCTAGTAGACACCATCTGGTTGAGTCAGTTTTTTTTATG  2557

GTGTATTTTGAACCCATTCTGATAGTCTCTTTTAACTGGAAGATTTCAATTACTTACGTTAATGTAATTATTAATATGT  2636

TAGGATTTATCCTCAGTCAGCCAGTTTGTTATGTCTTTTCTATTCTACTGTTATCACATTTGTACCACTTAAAGTGGAA  2715

TCTAGGCACTTTATCACCATTTAGATCCTATTACCTTTTCTCATCTAGGATATAGTTATCTTCTACATAATCTTTCTGT  2794

ATCTTAAAACCCATCAATAAATTATTATATATTTTCTACTTTTAATCACTCAGAAGATTTAAAAAACTCATGAGAAGAG  2873

TAATCTGTTATGTTTTTCCAGATATTTACCATTTCTGTTGCTCTTCCTTCATTATTTTCCAAATTTCGTTCTGCAAATT  2952

TCCACTTCTTCTGATAGACGTTTTTTAGTTCTTTTAGAGTGGTTCTGATAGGTACAGATTCTCTTATTTTTTGCTCCCT  3031

CTGAGGACATCTTTTTCTCACCTTCATTCTCAGTGATGTTTTTTGCTTGTAGTATTTTTAGTTGACATTGTTTTCTGTT  3110

CAGCAGTTTCCTTTTAGCTTCCGTATTTCCTGATGAGAAATCTGCAGTCATTCAAATTGTTGTTTCCCTGTATGTAGTG  3189

TGTCATTTTTCTGTCAGATTTCAAGGTATTTATCTTTAGTTTTTAGCCATTTCATTATGTTGGGGATGAGTTTCCTTGT  3268

TTTATTCCCTTTGGAATTTGCTCCAATTCATAAATTTGCAGTTTTATGTCTTTTACCAAACTTAGAGGTTTTCAGCCTA  3347

ATTTCTAAAAATACTTTTTATTAGCCTGATTTTCATCTTTATAGGAAATAGTTTAAGTGATGACAAGTTCCAATAGCTT  3426

ATATGCCCAGAAGGCCTTCAAAATAAGAATTTTGAAAGAATACAGAAAACAAACTTTTATATCCTTCTCATGTCTTCTA  3505

CTGTAAAATTCATATGCTTTGCTACTCTAAACCTAGTTTGAAATCAACAGTCTTGAGAATAGATGAAAATTTTGATGAA  3584

TAGTGGAATTCTTTTAAATGGAAACCTCTTACATGTGATTTTCCTTGCCATCTAGAAATAAACCATAGTATTTATGTTG  3663

AATCAATCAATATTATATTTTGTTTTTTTCCTCCTCTTCTGAGACTCTTATTGTGGAAATGTTAGACTTTTATGTTTTC  3742

CTAAATGTCCCTGATATTCTACTTATTTAGAACATCTTTTCATTTTTTCCATTATTCTGATTGGGTAATTTTAATTTGT  3821

CTATTTTCAAATTTGCTGGAGTGTTCACCTGTTGTTGTCTGTGTCGTCCCACTGAGTGCATTCACCACCTTTTAAATTT  3900

TGGTCACTGTATGTATCAGTTCTAAAATTTCCATTTTGTTCTCTATATTTTAAATTTCTTGGCTTATATTCTATTTTCC  3979
```

FIG. 4B

TGCAAATGTGTCAGCATTTGCTTGTTTGAGCTTTTTTTTTTTTCAAGACAGGGTCTCAACTCTGTTACCCAGGCTGGAGT 4058

GCAGTGGTGCGATCTCAGCTCACTGCAACCTCTGCCTCCTGGTTCAAGCGATTATTGTGCCTCAGCCTCCTGAGTAGCT 4137

GGGATTACAGGCATGCACCACCACAGCCCAGCTAATTTTTTGTATTTTTAGTAGAGACAGAGTTTTGCTATGTTGGCCA 4216

GGCTGGTTTTGAACTCCTGGCCTCAAGTGATCCACCCACCTCAGCCTCCCAAAGTGCTGGGATTACAGGCCACTACACC 4295

TGGCACATTTGAGTATTTTTTTTTTTTTTTTTTTTTTTTGAGATGGAGTCTCGCTCTGTCATCTAGGCTGGAGTGCAGTGG 4374

TGTGATCTCAGCTCACTGCAGCCTCTGTCTCCCGGGCTCAAGCGATTCTCTTGCCTCAGCCTCCTGAGTAGCTAGGACT 4453

ACAGGTGCATGCCAACACGCCCGGCTAATTTTTTTAAAAAATATTTTTAGTAGAGACAGGGTTTCACCATTTTGGCCAG 4532

GATGGTCTCGATCTCCTGACCTCATGATCCACCCGCCTCGGCCTTCCAAAGTGCTGGGATTACAGGCATGAGCCACCGT 4611

GCCTGGCCTCATTTGAGTATTTTTATAATGTCTCTTTTAAAGTCTTTGTCAGATAATTCCACTGTACATGTTATTCAGT 4690

GTTTGGTGTCCACTGAGTTGTCATTTGCCAGACAAGTGGAGATTTTTGCAGCTCATCCTTGTATTCTCAGTAGTTCCGA 4769

TATGTACCCTCGACATGTGAATGTTATCTTATGAGACTCTGTTTTATTTGTATCCAACAGAAGATGTTTATTATTTATT 4848

TGGCTTTCTGTGAACTGAGGTCTTAATATCAGCTCATTTTAAAAGTCTTTGCAGTGGTATTCGGATCTATCCTGTGTGT 4927

GCCTATGAGATTGGGTGCAGTGTATCCTGTTAGCTCCATTCTCAGGGCGTTTGAATGTGAATTAGGACCAGCGCAATGA 5006

ATGCTCAAGTTGGGGTTGGGCGTTAGAATTCATAAAAGTCTTTATATGCTCAG 5059

FIG. 4C

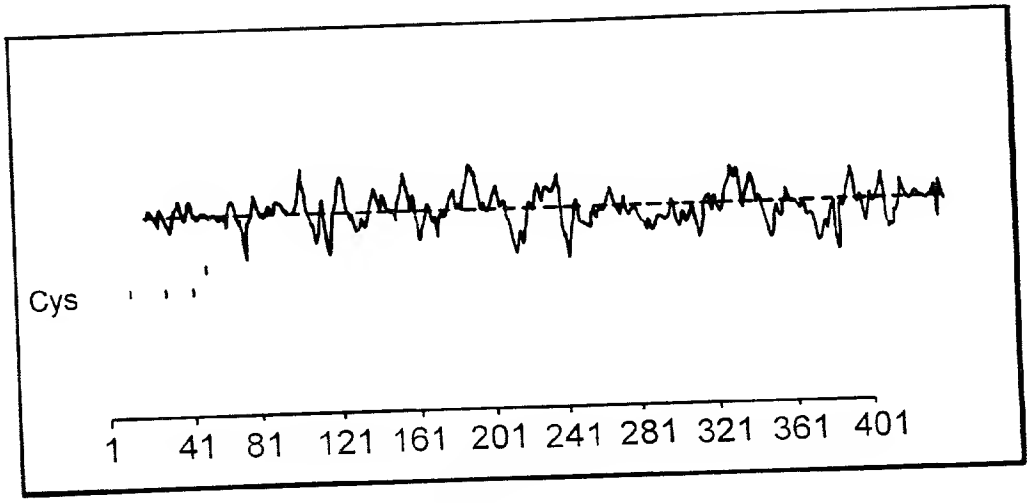

FIG. 5

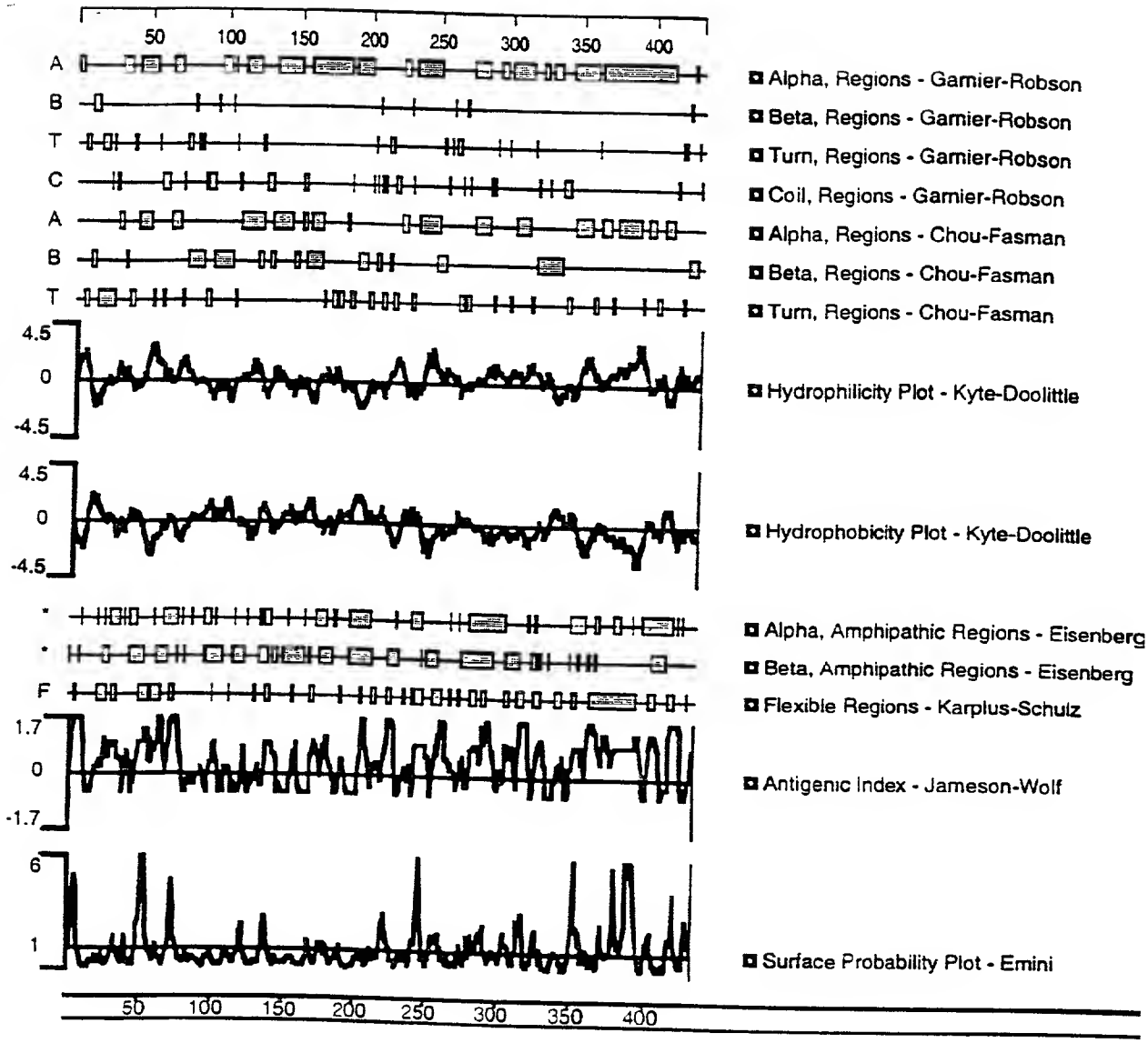
50
100
150
200
250
300
350
400
A
B
T
C
A
B
T
4.5
0
-4.5
4.5
0
-4.5
*
*
F
1.7
0
-1.7
6
1
50
100
150
200
250
300
350
400
Alpha, Regions - Garnier-Robson
Beta, Regions - Garnier-Robson
Turn, Regions - Garnier-Robson
Coil, Regions - Garnier-Robson
Alpha, Regions - Chou-Fasman
Beta, Regions - Chou-Fasman
Turn, Regions - Chou-Fasman
Hydrophilicity Plot - Kyte-Doolittle
Hydrophobicity Plot - Kyte-Doolittle
Alpha, Amphipathic Regions - Eisenberg
Beta, Amphipathic Regions - Eisenberg
Flexible Regions - Karplus-Schulz
Antigenic Index - Jameson-Wolf
Surface Probability Plot - Emini

FIG. 6

```
   1 CGCGTCCGGCTGCAGCGGGGTGAGCGGCGGCAGCGGCCGGGGATCCTGGAGCCATGGGGC
     GCGCAGGCCGACGTCGCCCCACTCGCCGCCGTCGCCGGCCCCTAGGACCTCGGTACCCCG
                                                    1▶ M   G

  61 GCGCGCGCGACGCCATCCTGGATGCGCTGGAGAACCTGACCGCCGAGGAGCTCAAGAAGT
     CGCGCGCGCTGCGGTAGGACCTACGCGACCTCTTGGACTGGCGGCTCCTCGAGTTCTTCA
   3▶R  A  R  D  A  I  L  D  A  L  E  N  L  T  A  E  E  L  K  K

 121 TCAAGCTGAAGCTGCTGTCGGTGCCGCTGCGCGAGGGCTACGGGCGCATCCCGCGGGGCG
     AGTTCGACTTCGACGACAGCCACGGCGACGCGCTCCCGATGCCCGCGTAGGGCGCCCCGC
  23▶F  K  L  K  L  L  S  V  P  L  R  E  G  Y  G  R  I  P  R  G

 181 CGCTGCTGTCCATGGACGCCTTGGACCTCACCGACAAGCTGGTCAGCTTCTACCTGGAGA
     GCGACGACAGGTACCTGCGGAACCTGGAGTGGCTGTTCGACCAGTCGAAGATGGACCTCT
  43▶A  L  L  S  M  D  A  L  D  L  T  D  K  L  V  S  F  Y  L  E

 241 CCTACGGCGCCGAGCTCACCGCTAACGTGCTGCGCGACATGGGCCTGCAGGAGATGGCCG
     GGATGCCGCGGCTCGAGTGGCGATTGCACGACGCGCTGTACCCGGACGTCCTCTACCGGC
  63▶T  Y  G  A  E  L  T  A  N  V  L  R  D  M  G  L  Q  E  M  A

 301 GGCAGCTGCAGGCGGCCACGCACCAGGGCTCTGGAGCCGCGCCAGCTGGGATCCAGGCCC
     CCGTCGACGTCCGCCGGTGCGTGGTCCCGAGACCTCGGCGCGGTCGACCCTAGGTCCGGG
  83▶G  Q  L  Q  A  A  T  H  Q  G  S  G  A  A  P  A  G  I  Q  A

 361 CTCCTCAGTCGGCAGCCAAGCCAGGCCTGCACTTTATAGACCAGCACCGGGCTGCGCTTA
     GAGGAGTCAGCCGTCGGTTCGGTCCGGACGTGAAATATCTGGTCGTGGCCCGACGCGAAT
 103▶P  P  Q  S  A  A  K  P  G  L  H  F  I  D  Q  H  R  A  A  L

 421 TCGCGAGGGTCACAAACGTTGAGTGGCTGCTGGATGCTCTGTACGGGAAGGTCCTGACGG
     AGCGCTCCCAGTGTTTGCAACTCACCGACGACCTACGAGACATGCCCTTCCAGGACTGCC
 123▶I  A  R  V  T  N  V  E  W  L  L  D  A  L  Y  G  K  V  L  T

 481 ATGAGCAGTACCAGGCAGTGCGGGCCGAGCCCACCAACCCAAGCAAGATGCGGAAGCTCT
     TACTCGTCATGGTCCGTCACGCCCGGCTCGGGTGGTTGGGTTCGTTCTACGCCTTCGAGA
 143▶D  E  Q  Y  Q  A  V  R  A  E  P  T  N  P  S  K  M  R  K  L

 541 TCAGTTTCACACCAGCCTGGAACTGGACCTGCAAGGACTTGCTCCTCCAGGCCCTAAGGG
     AGTCAAAGTGTGGTCGGACCTTGACCTGGACGTTCCTGAACGAGGAGGTCCGGGATTCCC
 163▶F  S  F  T  P  A  W  N  W  T  C  K  D  L  L  L  Q  A  L  R

 601 AGTCCCAGTCCTACCTGGTGGAGGACCTGGAGCGGAGCTGAGGCTCCTTCCCAGCAACAC
     TCAGGGTCAGGATGGACCACCTCCTGGACCTCGCCTCGACTCCGAGGAAGGGTCGTTGTG
 183▶E  S  Q  S  Y  L  V  E  D  L  E  R  S

 661 TCCGGTCAGCCCCTGGCAATCCCACCAAATCATCCTGAATCTGATCTTTTTATACACAAT
     AGGCCAGTCGGGGACCGTTAGGGTGGTTTAGTAGGACTTAGACTAGAAAAATATGTGTTA

 721 ATACGAAAAGCCAGCTTGAA
     TATGCTTTTCGGTCGAACTT
```

FIG. 7